



**University of  
Zurich<sup>UZH</sup>**

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2011

---

## **Class Evolution Tree: a graphical tool to support decisions on the number of classes in exploratory categorical latent variable modeling for rehabilitation research**

Kriston, Levente ; Melchior, Hanne ; Hergert, Anika ; Bergelt, Corinna ; Watzke, Birgit ; Schulz, Holger ; von Wolff, Alessa

**Abstract:** The aim of our study was to develop a graphical tool that can be used in addition to standard statistical criteria to support decisions on the number of classes in explorative categorical latent variable modeling for rehabilitation research. Data from two rehabilitation research projects were used. In the first study, a latent profile analysis was carried out in patients with cancer receiving an inpatient rehabilitation program to identify prototypical combinations of treatment elements. In the second study, growth mixture modeling was used to identify latent trajectory classes based on weekly symptom severity measurements during inpatient treatment of patients with mental disorders. A graphical tool, the Class Evolution Tree, was developed, and its central components were described. The Class Evolution Tree can be used in addition to statistical criteria to systematically address the issue of number of classes in explorative categorical latent variable modeling.

DOI: <https://doi.org/10.1097/MRR.0b013e3283460e7d>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-107236>

Journal Article

Published Version

Originally published at:

Kriston, Levente; Melchior, Hanne; Hergert, Anika; Bergelt, Corinna; Watzke, Birgit; Schulz, Holger; von Wolff, Alessa (2011). Class Evolution Tree: a graphical tool to support decisions on the number of classes in exploratory categorical latent variable modeling for rehabilitation research. *International Journal of Rehabilitation Research*, 34(2):181-185.

DOI: <https://doi.org/10.1097/MRR.0b013e3283460e7d>

# Class Evolution Tree: a graphical tool to support decisions on the number of classes in exploratory categorical latent variable modeling for rehabilitation research

Levente Kriston, Hanne Melchior, Anika Hergert, Corinna Bergelt, Birgit Watzke, Holger Schulz and Alessa von Wolff

The aim of our study was to develop a graphical tool that can be used in addition to standard statistical criteria to support decisions on the number of classes in explorative categorical latent variable modeling for rehabilitation research. Data from two rehabilitation research projects were used. In the first study, a latent profile analysis was carried out in patients with cancer receiving an inpatient rehabilitation program to identify prototypical combinations of treatment elements. In the second study, growth mixture modeling was used to identify latent trajectory classes based on weekly symptom severity measurements during inpatient treatment of patients with mental disorders. A graphical tool, the Class Evolution Tree, was developed, and its central components were described. The Class Evolution Tree can be used in addition to statistical criteria to systematically address the issue of number of classes in

explorative categorical latent variable modeling.

*International Journal of Rehabilitation Research*

34:181–185 © 2011 Wolters Kluwer Health | Lippincott Williams & Wilkins.

*International Journal of Rehabilitation Research* 2011, 34:181–185

**Keywords:** classification, latent class analysis, medical illustration, multivariate analysis, rehabilitation

Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Correspondence to Levente Kriston, PhD, Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, D-20246 Hamburg, Germany  
Tel: +49 40 7410 56849; fax: +49 40 7410 54965;  
e-mail: l.kriston@uke.uni-hamburg.de

Received 13 October 2010 Accepted 22 February 2011

## Introduction

Exploratory categorical latent variable modeling is increasingly used to identify not manifestly defined (latent) distinct classes in cross-sectional and longitudinal data. The goal of this approach is to classify individuals into groups (clusters) based on individual patterns of observed variables so that individuals within a certain group are more similar than individuals between groups (Marcoulides and Moustaki, 2002). In rehabilitation research, latent groups behind complex data patterns can be present in several situations, for example as a consequence of multifaceted supply structures (Vogel and Zdrahal-Urbaneck, 2004), various rehabilitation needs (Van Harten *et al.*, 1998; Fulton, 1999), received intervention blocks (Bürge *et al.*, 2008), or treatment response idiosyncrasies (Van Weert *et al.*, 2004). In addition to model specification issues (e.g. restriction of variance/covariance of observed variables within groups to be equal across groups or to a constant value), the central question of exploratory categorical latent variable modeling concerns the number of distinct classes.

Similar to other exploratory procedures (e.g. factor analysis, cluster analysis), several recommendations exist for deciding on the number of classes ( $k$ ) in categorical latent variable models (Nylund *et al.*, 2007). A straight forward practice is to use the Pearson  $\chi^2$  or the likelihood ratio goodness of fit test and choose the model with the lowest number of classes that yield acceptable fit (Goodman,

1974). However, validity of this approach is limited in cases when the expected frequencies for a large number of observed variable patterns are low (Huang, 2005). Instead of testing the goodness of fit of a specified model, it is usually preferred to use statistical information criteria for selecting among models with different numbers of classes. The most common information criteria are the Bayesian Information Criterion (Schwarz, 1978; Kass and Raftery, 1995) and Akaike's Information Criterion (Akaike, 1987). As these indexes depend on the number of estimated parameters (and thus, on model complexity), no universally valid threshold for an acceptable model fit can be established. Usually, several models with an increasing number of classes are tested and the model with the lowest value on these criteria is preferred. It is also possible to create a scree plot (Cattell, 1966) using the log-likelihood or the information criteria and to look for a 'substantive change' or an 'elbow' in the graph.

As an additional option to choose the most informative number of classes, a model with  $k$  classes can be directly tested against a model with  $k-1$  classes to see if the improvement by modeling one additional class is statistically significant. The most commonly used procedures include the Vuong–Lo–Mendell–Rubin likelihood ratio test (Vuong, 1989; Lo *et al.*, 2001) and the bootstrap likelihood ratio test (McLachlan and Peel, 2000). These tests can be used for example with an increasing number

of  $k$ , until the first statistically nonsignificant ( $P$  value of more than 0.05) comparison occurs, in which case the model with  $k-1$  classes in this comparison should be preferred. Especially the bootstrap test showed promising characteristics in numerous models (Nylund *et al.*, 2007). A novel model selection algorithm operating also on the clustering variables in addition to the number of classes is described by Dean and Raftery (2010). Further commonly used criteria to decide on the number of classes include requiring high accuracy of classification (entropy) in a model (near 1.0, e.g. more than 0.80) and not less than a certain proportion (e.g. 1, 5%) of total count in a class.

However, most experts agree that determining the number of classes cannot depend solely on statistical measures. One should also consider other factors including the research objective, theoretical background, interpretability, and practical applicability (Muthén and Muthén, 2000; Bauer and Curran, 2003; Jung and Wickrama, 2008). The aim of our study was to develop a graphical tool that addresses these aspects and can be used in addition to statistical criteria to support decisions on the number of classes in explorative categorical latent variable modeling.

## Methods

Two doctoral dissertation projects at the Department of Medical Psychology of the University Medical Center Hamburg-Eppendorf (Hamburg, Germany) aimed to identify distinct classes of patients treated in inpatient rehabilitation clinics by using exploratory categorical latent variable modeling.

The first study (study A) focused on the identification of prototypical combinations of treatment elements in the inpatient rehabilitation of 678 patients with breast cancer. Information on such treatment combinations would help clinicians to base their decisions concerning the allocation of interventions to patients on empirical findings. In this cross-sectional study, the received rehabilitation treatment was described by 12 variables, each of them recording the duration (in minutes) of specific treatments (e.g. physiotherapy, psychotherapy, psychoeducation) during the inpatient stay. This approach with a categorical latent variable and metric observed variables is frequently termed as latent profile analysis. The researchers assumed that prototypical combinations of treatment elements exist that may describe the provided care in a more informative and parsimonious way than treatment profiles for every single patient. Covariance of variables within classes was restricted to zero and models with an increasing number of classes were tested. In preliminary analyses statistical criteria suggested solutions between five and seven classes.

The second study (study B) aimed to identify prototypical trajectories of symptom change in inpatient rehabilitation treatment of 576 patients with mental disorders. This is judged essentially to predict treatment short-term and long-term outcomes already during

treatment. In this longitudinal study, data on symptom severity were collected weekly over 8 weeks. The longitudinal (growth) model included an intercept, a linear slope, and a quadratic slope, which were allowed to covariate within classes. This approach with correlated growth parameters is usually described as growth mixture modeling. The researchers hypothesized that a limited number of distinct classes of symptom trajectories exist that allow for a categorization of patients according to their symptom courses. In preliminary analyses statistical criteria suggested solutions between four and seven classes.

We first used standard statistical criteria to describe the tested models while using latent class modeling techniques in these studies. However, we experienced several problems with this 'math-only' approach. First, most indexes allow only for model comparisons rather than assessment of absolute fit. Second, the criteria disagreed to a substantial degree suggesting different solutions. Third, similarity (stability) of the solutions across models could not be judged, due to availability of information only on single models. And fourth, no content-related criteria could be included in the decision-making process concerning the choice between models. To overcome these shortcomings, we were looking for an alternative way of rendering our analyses. We have found a graphical illustration of the process of fitting models with a varying number of classes extremely helpful to support decision on the number of classes. We named this tool the Class Evolution Tree (CET).

## Results

The CET is similar to a hierarchical cluster tree, or dendrogram, in hierarchical cluster analysis. The CET displays all models in one graph starting with the solution with one class at the top, a model with two classes one level below, a model with three classes further below, and so on. For each class in a certain model with  $k$  classes, the frequency distribution of the participants according to the categorization in the model with  $k+1$  classes is displayed (descending or target proportions). In addition, the frequency distribution of the participants according to the categorization in the model with  $k-1$  classes is shown (ascending or source proportions). Optimally starting from the top, each class in each model should be named according to its characteristics (i.e. pattern of the observed variables as presented in standard statistical output), thus enabling the analysts to interpret all models occurring during the search for the best model. In this way, it becomes observable which classes stay stable through the models with different number of classes, how they are being split, become more specific or general, disappear or maybe reappear again on a lower level, etc. In general, it becomes visible how the identified classes evolve.

Figure 1 shows the produced CETs for our two studies. For both CETs, each solution is presented in a separate row. Classes are represented by boxes with global

descriptive information: class label, number of cases in the class, and fraction of cases in the class. Similarity of classes across solutions is represented by lines connecting classes (boxes) across neighboring levels of the CET. For our studies, we defined substantial similarity (significant case flow) as an overlap between classes at which either the descending or the target proportion exceeds 20%. Accordingly, we suppressed all numerical information on these proportions that fell below 20%.

We identified several interesting phenomena in our analyses. Some classes remained stable across solutions

with the number of cases varying only marginally (e.g. class 'Psychotherapy and Education' in study A). On the contrary, other classes emerged, disappeared in a solution with a higher number of classes, but reappeared again when the number of classes was further increased (e.g. class 'Physiotherapy and recreation' in study A). In addition, classes could be identified that appeared comparably late (i.e. in a solution with a high number of classes) and unexpected without being present in any of the models with fewer classes (e.g. class 'Early response' in study B). Furthermore, we could identify cases that were 'oscillating' between classes across the tested models complicating a

**Fig. 1**

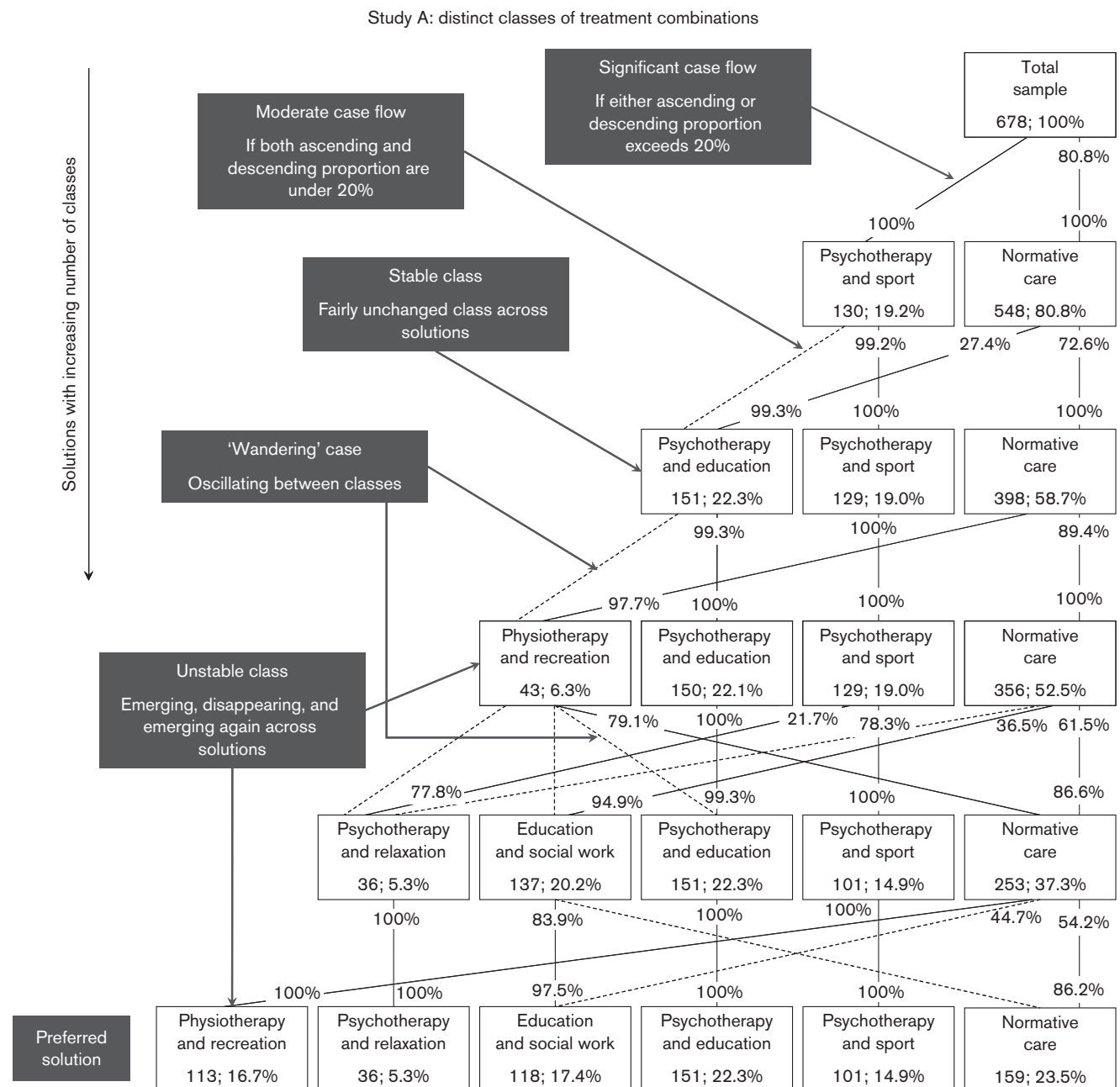
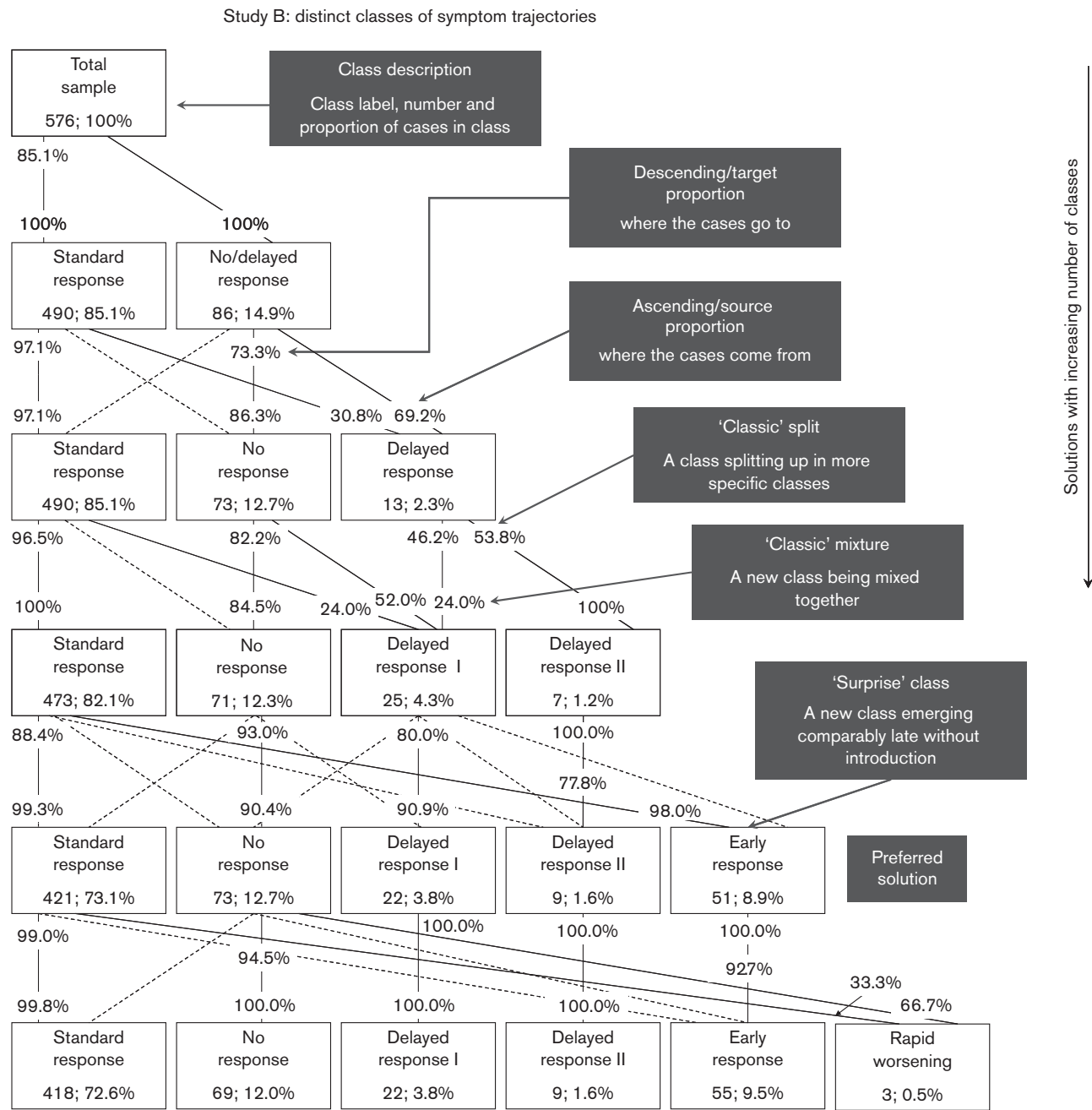


Fig. 1 (continued)



Class Evolution Tree in two studies.

clear allocation. We could also observe a theoretically supported split of a certain class in more specific classes (e.g. class 'Delayed response' in study B) and the emergence of new classes as a mixture of several others (e.g. class 'Delayed response I' in study B).

In both studies, we included the CETs in the decision-making with regard to the most informative number of classes. After limiting the number of possible solutions using

statistical criteria, the CETs were used to decide which models are theoretically and practically the most comprehensible and interpretable and should therefore be preferred.

### Discussion

The CET proved to be a helpful instrument to support decision on the number of classes in exploratory categorical latent variable modeling for rehabilitation research.

A possible limitation of the tool is the rapidly increasing complexity with increasing number of classes. However, it is usual in rehabilitation research to search for a relatively low number of distinct classes, because the information gain of the categorizing approach is quite low when the number of classes exceeds the number of observed variables and model complexity is limited by the sample sizes between 200 and 2000 participants that are common in this setting.

An advantage of the CET is its flexibility. Analysts are free to suppress or add information according to their preferences. For example, in the presented examples proportions less than 20% are not displayed. It is a convenient decision that can be modified according to the research topic. Furthermore, in models with a limited number of classes it is also possible to examine overlaps between classes in models differing in more than one level.

Although the CET is a visual tool, the development of statistical indices describing the models (or parts of them) may be promising. For example, a weighted index of descending and ascending probabilities could adequately describe the 'stability' of a certain model or a specific class. Similarly, averaging the highest-class membership probabilities for each class and comparing the averages between levels could be informative of the 'uniqueness' of the identified groups. In general, we encourage statisticians to develop indexes that describe models with  $k$  classes in relation to models with  $k-1$  and  $k+1$  classes rather than 'stand-alone' solutions.

Furthermore, CETs can also be created not only for latent class models but also for any nonhierarchical classification procedure, such as  $Q$  factor analysis and  $k$ -means clustering. As the choice of a certain solution with a definite number of groups is also central in these techniques, they could probably profit from the application of the CET visualization. In this context, the nonignorable advantage of unifying decision criteria across different statistical techniques should also be highlighted.

We also think that the creation of a CET could easily be implemented in software packages that estimate categorical latent class models. Apart from the class labels, which are imputed by the researcher, all information displayed in a CET is determined by the estimated models. Although the sequence of the classes in each row (model) is somewhat arbitrary, it can easily be standardized, for example through presenting classes from left to right with a decreasing number of cases. When implemented

in standard software packages, this tool would become available for a large group of researchers.

## Acknowledgements

The authors thank the visitors of the IV European Congress of Methodology (July 21–23, Potsdam, Germany) for their helpful comments on a poster presentation of the topic. This study was not funded externally. The studies that were used to show applications were funded by the Wilhelm-Stiftung für Rehabilitationsforschung (Wilhelm Foundation for Rehabilitation Research, Germany).

None of the authors report conflicts of interest.

## References

- Akaike H (1987). Factor analysis and AIC. *Psychometrika* **52**:317–332.
- Bauer DJ, Curran PJ (2003). Distributional assumptions of growth mixture models: implications for overextraction of latent trajectory classes. *Psychol Methods* **8**:338–363.
- Bürge E, Cieza A, Allet L, Finger ME, Stucki G, Huber EO (2008). Intervention categories for physiotherapists treating patients with internal medicine conditions on the basis of the International Classification of Functioning Disability and Health. *Int J Rehabil Res* **31**:43–50.
- Cattell R (1966). The scree test for the number of factors. *Multivariate Behav Res* **1**:245–276.
- Dean N, Raftery AE (2010). Latent class analysis variable selection. *Ann Inst Stat Math* **62**:11–35.
- Fulton C (1999). Patients with metastatic breast cancer: their physical and psychological rehabilitation needs. *Int J Rehabil Res* **22**:291–301.
- Goodman LA (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* **61**:215–231.
- Huang G (2005). Selecting the number of classes under latent class regression: a factor analytic analogue. *Psychometrika* **70**:325–345.
- Jung T, Wickrama KAS (2008). An introduction to latent class growth analysis and growth mixture modeling. *Soc Personal Psychol Compass* **2**:302–317.
- Kass RE, Raftery AE (1995). Bayes factors. *J Am Stat Assoc* **90**:773–795.
- Lo Y, Mendell NR, Rubin DB (2001). Testing the number of components in a normal mixture. *Biometrika* **88**:767–778.
- Marcoulides GA, Moustaki I (2002). *Latent variable and latent structure models*. Mahwah, NJ: Lawrence Erlbaum Associates.
- McLachlan G, Peel D (2000). *Finite mixture models*. New York: Wiley.
- Muthén BO, Muthén LK (2000). Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. *Alcohol Clin Exp Res* **24**:882–891.
- Nylund KL, Asparouhov T, Muthén BO (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. *Struct Equ Modeling* **14**:535–569.
- Schwarz G (1978). Estimating the dimension of a model. *Ann Stat* **6**:461–464.
- Van Harten WH, Van Noort O, Warmerdam R, Hendricks H, Seidel E (1998). Assessment of rehabilitation needs in cancer patients. *Int J Rehabil Res* **21**:247–257.
- Van Weert E, Hoekstra-Weebers JEHM, Grol BMF, Otter R, Arendzen JH, Postema K, van der Schans (2004). Physical functioning and quality of life after cancer rehabilitation. *Int J Rehabil Res* **27**:27–35.
- Vogel H, Zdrahal-Urbaneck J (2004). Rehabilitation in Germany: new challenges of a structured social security scheme. *Int J Rehabil Res* **27**:93–98.
- Vuong QH (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* **57**:307–333.